DOCUMENT RESUME

ED 231 146 EC 152 523

AUTHOR Horn, Carin E.; Scott, Brian L.
TITLE Micro-Based Speech Recognition: Instructional Innovation for Handicapped Learners.
PUB DATE Mar 83
NOTE 18p.; Paper presented at the Council for Exceptional Children National Conference on the Use of Microcomputers in Special Education (Hartford, CT, March 10-12, 1983).
PUB TYPE Viewpoints (120) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Accessibility (for Disabled); *Articulation Impairments; *Computer Assisted Instruction; *Input Output Devices; Man Machine Systems; *Microcomputers; *Physical Disabilities; Pronunciation; *Speech
IDENTIFIERS *Voice Based Learning System

ABSTRACT

A new voice based learning system (VBLS), which allows the handicapped user to interact with a microcomputer by voice commands, is described. Speech or voice recognition is the computerized process of identifying a spoken word or phrase, including those resulting from speech impediments. This new technology is helpful to the severely physically handicapped individual who cannot use a keyboard. The VBLS makes many previously inaccessible electronic instructional aids available to physically disabled learners, and involves them in their own educational/vocational progress. The VBLS system's software is written in assembly code and BASIC. During the study or learning component of the system, the student user says a lesson vocabulary word/phrase, reads the related question, and vocally answers the question aloud into the system's microphone by saying the prompted correct response. Master templates are adapted to each user's unique speech patterns as part of the instructional design logic. The VBLS system can be adapted to varying skill levels: physical, cognitive, and communicative. Additionally, a word drill study option provides the student with real time pronunciation feedback. Finally, there are over 40 options available so that a lesson can easily be adapted to focus on the student user's ability and learning sequence preferences. Information is presented on: the logic of a speech recognition system, and configuration of an environmental control system the algorithm for word verification, the instrumentation and cost, and general principles of how the system works. (SEW)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED231146

# MICRO-BASED SPEECH RECOGNITION: INSTRUCTIONAL INNOVATION FOR HANDICAPPED LEARNERS

Carin E. Horn

Brian L. Scott

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Carin E. Horn
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

There is an apparent need to develop and implement policy to guide the application of information and communications technologies to all levels of education. Full advantages of the socially desirable educational opportunities offered by a broad range of technologies must be pursued to increase intellectual productivity and enhance the quality of life.

Hon. George E. Brown, Jr., 1979[1]

## Voice Input and the Handicapped Learner

With over 20 million handicapped persons in the United States, there is no one user-profile describing the special needs learner. The most common characteristic shared by disabled persons, however, is the difficulty, and often inability, to communicate with other people in a meaningful way. The need for basic communication skills, which are highly personal and individualized, becomes intensified when the range of physical functions decreases. Physically handicapped individuals have, in the past, needed a range of motion to access a computer's keyboard (or an extended keyboard) using a finger, a mouthstick, headstick, toe, or

EC152523

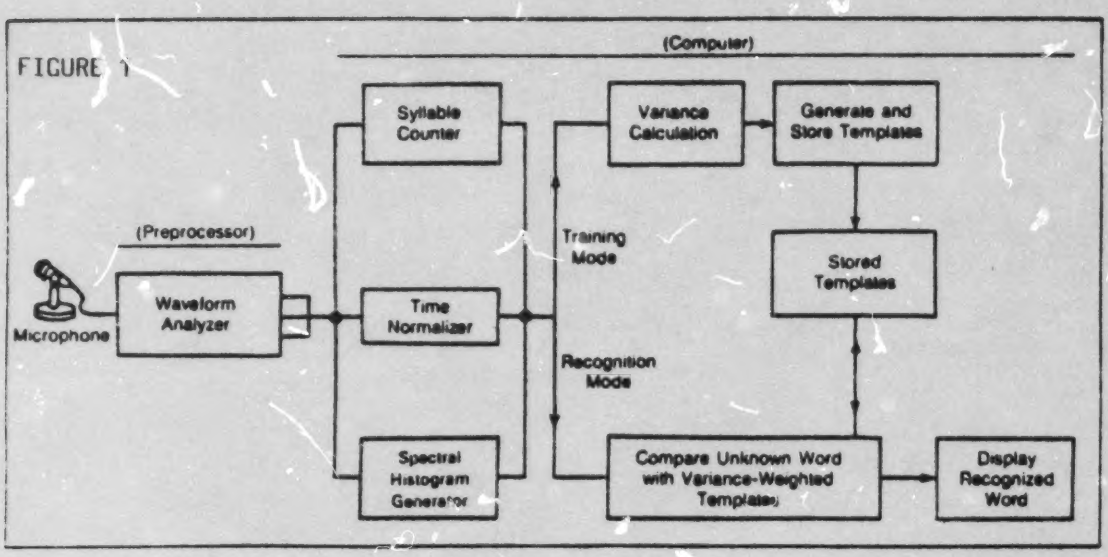
so on. Effective computing, whether for communication enhancement, environmental control, or instructional purposes, was dependent upon the speed and accuracy of user motor skills. All expressive thoughts were translated into keyboard manipulations of some description.

The VBLS™ voice-based learning system, which was designed for able-bodied and physically disabled student-users, may be run with no user keyboard entry whatsoever. The VBLS system, which will be reviewed in this article, is a low cost, speech-controlled authoring system. The System takes drill and repetition learning beyond relative linguistic competence, The VBLS word verification algorithm for speech recognition encourages communicative competency within a specific context, regardless of reading level or speech disorder.

## Speech Recognition

Speech recognition (voice recognition) is the computerized process of identifying a spoken word or phrase by matching an unknown utterance to a number of known utterances and selecting the closest match within prescribed tolerances. The basic logic of a speech recognition system involves processing an acoustic waveform with a filter bank, an LPC (Linear Predictive Coding) analysis, or zero-crossing analysis to extract spectral data. Templates of the resulting digital data are then stored and become the active vocabulary or language model for the spoken words or phrases.

Recognition itself occurs when a search and match is made between newly input speech and existing vocabulary templates. The matching or mapping process varies with the recognition algorithm being used. Figure 1 illustrates the schematic logic of a speech recognition system.




Speech recognition technology can be cost-effectively used in numerous applications, including data entry, hands-and-eyes busy tasks, distant or mobile inputs, data verification, or when data perishability is a problem. Voice entry terminals enable severely physically handicapped individuals to operate a computer by speaking to it in any language. Applications specific to disabled persons are most frequently expressive communication aids or

environmental control systems, wherein, the handicapped user directs several microcomputer-based electronic devices by voice commands. (See Figure 2.) Aside from facilitating environmental control, speech recognition technology is providing innovative educational opportunities for the physically handicapped. Voice makes many electronic instructional aids, previously inaccessible to the physically disabled, available for special education. Such voice-controlled devices actively involve the handicapped learner in his or her own educational/vocational progress.

Figure 2

Basic Configuration of an
Environmental Control System

There are a variety of Apple peripherals and assorted devices that may be used to create a voice-based environmental control system for handicapped users. The system components listed below are appropriate for individuals with high level spinal cord injury, muscular dystrophy, amyotrophic lateral sclerosis or other similar conditions which prohibit standard operation of a computer keyboard.

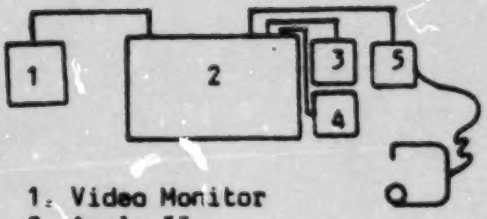
*Apple-compatible microcomputer (48K) with dual drives
*VETtm series voice entry terminal
Catnovation with DCH chip
Applecat BSR module
Applecat expansion module

or

A Thunderclock with a Hayes micro-modem
BSR ultrasonic transducer
BSR controller

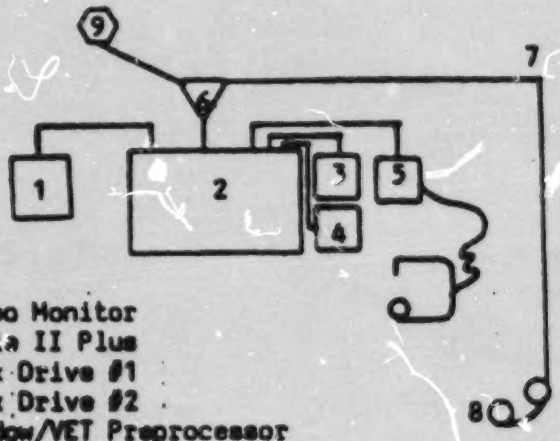
and lamp and/or appliance modules as needed

## CONFIGURATION 1



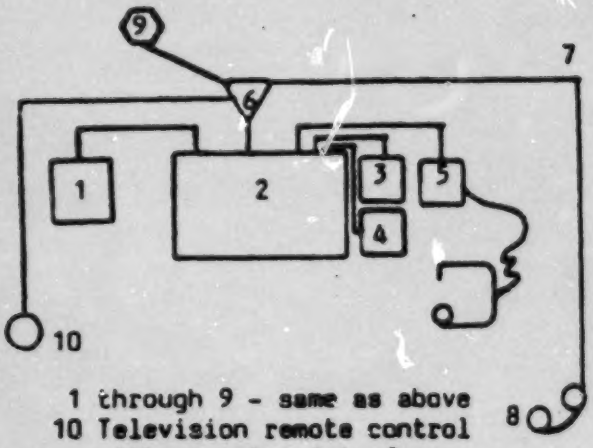

1. Video Monitor
2. Apple II
3. Disk Drive #1
4. Disk Drive #2
5. Shadow/VET Preprocessor with microphone

## CONFIGURATION 2



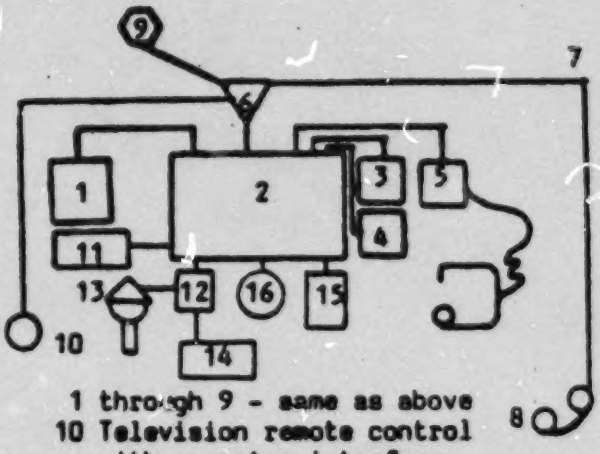

1. Video Monitor
2. Apple II Plus
3. Disk Drive #1
4. Disk Drive #2
5. Shadow/VET Preprocessor with microphone
6. Novation AppleCAT Expansion Module
7. Phone Line
8. Telephone Headset
9. Novation A/C Control

## CONFIGURATION 3




1 through 9 - same as above
10 Television remote control with computer interface

## THE C2E2 SYSTEM




1 through 9 - same as above
10 Television remote control with computer interface
11 Speaker
12 Power control
13 Lamp
14 Radio
15 Printer
16 Switches for Push Button Control

## THE VBLS VOICE-BASED LEARNING SYSTEM

### Research and Development Activities

Task: Develop voice-based learning system for non-computer oriented users in which voice input and system acceptance of same are essentially instantaneous.

Objective: The objective is to research, develop, design, and market a low cost, speech-controlled authoring system to meet educational needs of physically handicapped and able-bodied learners that will reinforce learning through pronunciation of correct responses. The resulting courseware would enable non-biased, non-discriminatory tutoring, reviewing, and testing of handicapped and non-handicapped student populations.

Rationale/Pronunciation and Meaningfulness: The instructional technology of the VBLS system is a systematic method for designing educational materials that incorporate an innovative communication medium -- speech recognition. The relationship between pronunciation and meaningful learning is significant because pronunciation serves to decode the written word into its oral form which can then elicit a meaningful response, if in fact, the word is in the oral vocabulary of the learner.[2] In this sense, speech becomes the bridge between conscious awareness and largely unconscious cognitive operations, such as identification,

classification, and storage.[3] The ability to read is basically dependent upon the skill relationships between visually displayed words and the spoken language which the learner already possesses.[4]

In an extensive examination of meaningfulness, Underwood and Schultz found that pronounciability was frequently an excellent predictor of learning. In one study using paired-associate procedures (Experiment 11), for thirty-six items obtained from three separately learned lists, the correlation between frequency and mean number of correct responses over fifteen trials was .09. On the other hand, the correlation for pronounciability and mean number correct was .57. In another experiment (Experiment 12), the correlation between pronounciability and correct responses was .76.[5] DiVesta and Ingersoll (1969) also found that paired-associate learning was related to ease of pronounciability.[6]

In summary, learning is reinforced through the pronunciation of contextually meaningful utterances.

Impact: The intention is that any non-programming educator/trainer could author highly individualized, speech-controlled educational materials within a structure flexible enough to support instructional designs in any discipline and within any reading level. The resulting courseware would evolve into a library of voice-based instructional materials.

Instructional Design Team: William D. Wagers, Systems Development/Systems Analyst; Carin E. Horn, Instructional Systems/ Educator; Warren C. Jones, Systems Development/ Programmer, and Janice D. Drake, Instructional Systems/ Communications Liasion.

Product Development to Date: The VBLS voice-based learning system prototype was completed in June 1982. It was presented at the National Educational Computing Conference in Kansas City, Missouri. The VBLS system is comprised of an Apple-compatible voice entry terminal by Scott Instruments with a Shure SM-10B or equivalent noise-cancelling microphone, a reference manual, and VBLS software on a 5¼ inch floppy disk. The voice entry terminal measures approximately 1¾" H x 6" W x 9" D and weighs three pounds. Power consumption is between 1-10w. Suggested retail price is $895.00, F.O.B., Denton, Texas.

The prototype was field tested in a 'typical' first grade public school classroom (Spring 1982) and has been subsequently critiqued by numerous professionals in various disciplines, including special education. The critiques have resulted in modifications to the original design, and are incorporated into the most current version of the VBLS system. Further peripheral options and adaptations are anticipated to highly customize the VBLS voice-based learning system for special needs learners.

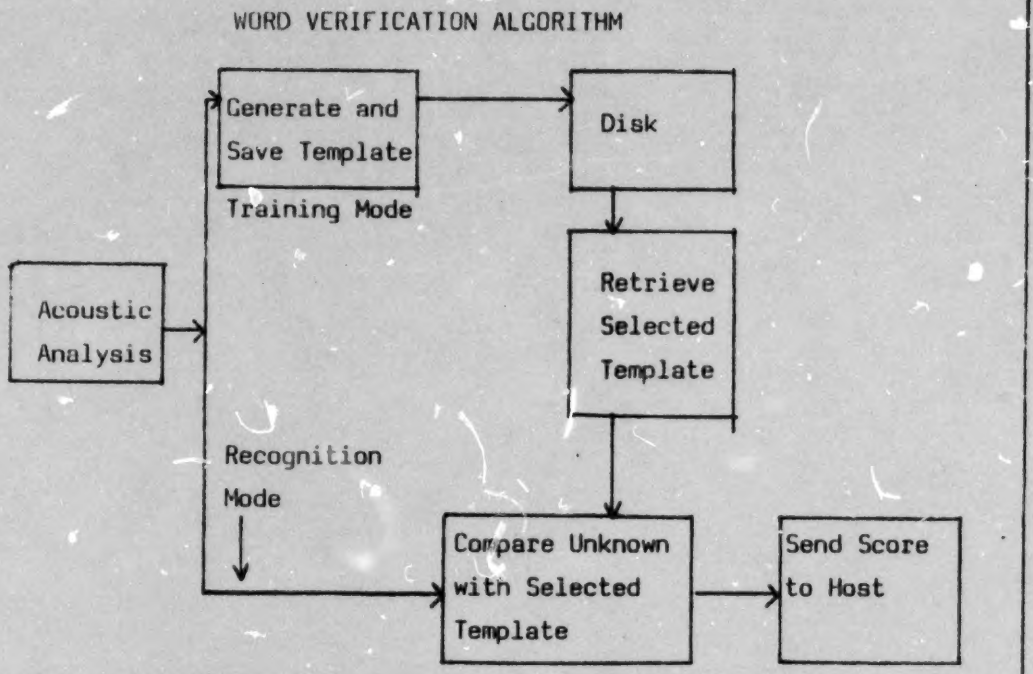
System Overview: The VET-2™ voice entry terminal emulates the microcomputer's keyboard, allowing the user to voice control the computer's microprocessor. This preprocessor accepts human utterances in the form of isolated words/phrases, including those resulting from speech impediments. Input sound waves are converted to digital data, and several samples (usually five) of the same word/phrase are averaged to construct a master template in the computer's memory which is then stored onto disk. [The Shadow/VET™, another Scott voice entry terminal, has 16K onboard memory and does not use the host computer's memory to store templates.] The speech analysis technique involves amplitude envelope and zero-crossing detection within the frequency regions of 250-1000Hz and 1000-5000Hz. Templates are derived through linear time compression, syllable counting, spectral histograms, and variance measurements of the data. Specification of active vocabulary, the effective vocabulary, and the total vocabulary size are determined by software control. The VBLS terminal is not intended for speaker-independent use and training is required.

The VBLS systems software is written in assembler code and BASIC. There are approximately 2600 lines of BASIC and 3200 lines of assembly language source in the VBLS authoring system.

User's Perspective: The process of training (enrolling) the preprocessor to recognize a specific word or phrase is, essentially, invisible to the VBLS end-user. The technique of word verification* coupled with the VBLS tutorial structure camouflages speaker-adaptive procedures. Voice is an integral component of the teaching/learning system, and training becomes part of that process, for example:

*[NOTE: Word verification is an algorithm that employs the concept of the one-word/one-phrase vocabulary. Regardless of the total vocabulary size, only one word or phrase is acceptable at any point in time. The recognition task is to verify the active word, not to select the correct word from among several candidates. As a result, more data can be kept on each word and more processing time can be devoted to the single word/phrase. The vocabulary becomes completely disk-based and is limited in size only by the mass storage capabilities of the system being used. An Apple floppy disk can store a sixteen hundred (1,600) word vocabulary. (See Figure 3.)]

FIGURE 3




User's Perspective, continued:

During the lesson authoring, the System randomly prompts a training sequence for tokens of words, continuing to do so until N passes of a word/phrase have been made. Training is a pre-set sequence, wherein, the author types the alphanumeric word/phrase identity into the keyboard terminal, then also speaks the same when prompted. ~~(See Figure 4.)~~

During the Study Session, the student-user is prompted to say a lesson vocabulary word/phrase, to read the related question, and to vocally answer the question aloud into the System's microphone by saying the prompted correct response. Master templates are, hereby, adapted to each user's unique

speech patterns as part of the instructional design logic. Once completed, the student does not retrain author-made vocabulary for that lesson, but accesses the adapted templates on disk to be recognized by the V8LS system in the Review. Adaptive procedures exist to update templates, if necessary. ~~(See Figure 5.)~~

FIGURE 4

LESSON AUTHORING

½ tone b/w photo series of display screens (being processed)

FIGURE 5

STUDY SESSION

½ tone b/w photo series of display screens (being processed)

Typical Use: The VBLS system is an authoring structure and instructional delivery system. The relative value of the resulting courseware is dependent upon the quality of instruction that the System delivers. Materials may be presented in small, easily understood segments or as a series of comprehensive questions and answers that also include supplemental information. The VBLS system is flexible enough to adapt to varying skill levels -- physical, cognitive, and communicative. Lesson practicality and originality are the author's responsibility.

The VBLS special needs learner should have some visual acuity, minimal reading skills (K-1), and the ability to help assign meaningful alphanumeric identifiers to spoken language components. For example, if a student refers to 'water' as "wa" and consistantly uses this utterance in a water-related way, the VBLS system can be trained to recognize "wa" to mean 'water.' Once oral language vocabulary components have been identified, lessons may be authored to cultivate communication competency using that individual student's uniquely spoken language.

VBLS lessons may be authored for instruction in the home as well as in the classroom. A unit may be designed for a specific population or an individual student. The menu-driven authoring format enables the teacher/parent to input whatever materials are student-appropriate. ~~(See Figure 6.)~~ The VBLS system will support unlimited types of curriculum. System

conventions require that all spoken responses be said within three seconds. The VBLS system 'assumes' that the author has had no technical training, save in his/her own area of expertise. No programming skills are needed to author VBLS courseware. No keyboard strokes are needed to study VBLS lessons. This authoring system option eliminates the need for keyboard-related motor movements. A typical Study Session would include voice-based tutoring, drill and repetition, review, and testing (optional).

WORD DRILL, which is another author-designated, Study Session option, provides the student with real time pronunciation feedback. Lesson vocabularies may be practiced at the word, phrase, or short sentence level within the three seconds convention. This option is particularly useful when improved student articulation is an instructional goal. WORD DRILL displays each of the correct responses for a particular lesson to the student who is then directed to say each into the microphone. The System compares the student's pronunciation to the author's template and displays a pronunciation score. Scores will range from 0 to 1000, with 1000 representing a perfect match. This technique allows the learner to progressively refine his/her pronunciation. Vocabulary words/phrases may be repeated as many times as the student chooses. Progressing to the next word currently requires two keyboard strokes (Space Bar and RETURN).

The VBLS system allows for built-in overlearning and diversified teaching/learning styles. There are over forty options available so that any one lesson can easily be adapted to focus on each student-user's ability and learning sequence preferences. One student, for example, can review the lesson vocabulary, take a pre-test, be tutored, reviewed, and tested on a lesson. Another might be instructor-directed to use WORD DRILL before running the tutorial and its review. A third student might complete the tutorial and then run its review three times. Each student Study Disk may be a highly customized version of the same exact tutorial information. The VBLS system provides the opportunity to develop voice-controlled learning materials for handicapped and non-handicapped student populations. VBLS lessons are adapted to recognize each individual student's system of oral communication.

## Comments

The human voice is interfacing instructional innovation and the physically handicapped learner. In the past, the absence of quality education for handicapped children necessitated Public Law 94-142, the Education for All Handicapped Children Act. Today, micro-based speech recognition is helping to implement the law, and the VBLS system is porviding the structure to author non-biased, non-discriminatory instructional courseware for handicapped and able-bodied learners.

## REFERENCES


1 Hon. George E. Brown, Jr., U.S. House of Representatives Committee on Science & Technology. Information and Communications Technologies Appropriate in Education (Including H.R. 4326) (Hearing). Washington, D.C.: U.S. Government Printing Office, 1979, p. 1.

2 Elizabeth S. Ghatala, Joel R. Levin, & Larry Wilder. Pronunciation & the Frequency Meaningfulness Effect in Children's Frequency Discrimination (Technical Report No. 344). Madison, Wisconsin: Wisconsin Research & Development Center for Cognitive Learning, May 1975, p. 7.

3 McNeill. The Two-fold Way for Speech (Paper) cited in J.S. Bruner, Language as an Instrument of Thought in Alan David (Ed.), Problems of Language and Learning. London: Heinemann, 1975, p. 79.

4 Margaret M. Clark. Language & Reading: Research Trends in Alan David (Ed.), Problems of Language and Learning. London: Heinemann, 1975, p. 91.

5&6 John F. Hall. Verbal Learning and Retention. Philadelphia: J.B. Lippincott, 1971, pp. 102-103.